Accounting for model error in variational data assimilation

A.S. Lawless

School of Mathematical and Physical Sciences
University of Reading

Work supported by NERC

Outline

Four-dimensional variational assimilation (4D-Var)

The 4D-Var data assimilation problem can be expressed as the minimization of

$$\mathcal{J}[\mathbf{x}_0] = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}^b)^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}^b) + \frac{1}{2}\sum_{i=0}^{N}(\mathcal{H}_i[\mathbf{x}_i] - \mathbf{y}_i)^{\mathrm{T}}\mathbf{R}_i^{-1}(\mathcal{H}_i[\mathbf{x}_i] - \mathbf{y}_i)$$

subject to the dynamical system

$$\mathbf{x}_{i+1} = \mathcal{M}_i(\mathbf{x}_i)$$

where

$\mathbf{x}^b$	A priori (background) estimate
$\mathbf{y}_i$	Observation
$\mathbf{B}$	Background error covariance matrix
$\mathbf{R}_i$	Observation error covariance matrix
$\mathcal{H}_i$	Observation operator

Incremental 4D-Var (Gauss-Newton)

We usually solve the 4D-Var problem by a series of linear quadratic minimizations of the form

$$\begin{aligned}
\tilde{\mathcal{J}}^{(k)}[\delta\mathbf{x}_0^{(k)}] &= \frac{1}{2}(\delta\mathbf{x}_0^{(k)} - [\mathbf{x}^b - \mathbf{x}_0{}^{(k)}])^{\mathrm{T}}\mathbf{B}^{-1}(\delta\mathbf{x}_0^{(k)} - [\mathbf{x}^b - \mathbf{x}_0{}^{(k)}]) \\
&+ \frac{1}{2}\sum_{i=0}^{N}(\mathbf{H}_i\delta\mathbf{x}_i^{(k)} - \mathbf{d}_i^{(k)})^{\mathrm{T}}\mathbf{R}_i^{-1}(\mathbf{H}_i\delta\mathbf{x}_i^{(k)} - \mathbf{d}_i^{(k)})
\end{aligned}$$

with

$$\begin{aligned}
\mathbf{d}_i &= \mathbf{y}_i - \mathcal{H}_i[\mathbf{x}_i^{(k)}] \\
\delta\mathbf{x}_{i+1} &= \mathbf{M}_i\delta\mathbf{x}_i
\end{aligned}$$

This is equivalent to a Gauss-Newton iteration (Lawless *et al.* (2005), Quart. J. Roy. Met. Soc.).

Alternative notation

For the linear case we can write the cost function as follows:

$$\mathcal{J}(\mathbf{x}_0) = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}^b)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}^b) + \frac{1}{2}(\hat{\mathbf{y}} - \hat{\mathbf{H}}\mathbf{x}_0)^T \hat{\mathbf{R}}^{-1}(\hat{\mathbf{y}} - \hat{\mathbf{H}}\mathbf{x}_0),$$

$$\hat{\mathbf{y}} = \begin{pmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \\ \vdots \\ \vdots \\ \vdots \\ \mathbf{y}_N \end{pmatrix} \quad \hat{\mathbf{H}} = \begin{pmatrix} \mathbf{H}_0 \\ \mathbf{H}_1\mathbf{M}_{0\to 1} \\ \vdots \\ \vdots \\ \vdots \\ \mathbf{H}_N\mathbf{M}_{0\to N} \end{pmatrix} \text{ and } \hat{\mathbf{R}} = \begin{pmatrix} \mathbf{R}_0 & 0 & \cdots & \cdots & 0 \\ 0 & \mathbf{R}_1 & 0 & \cdots & 0 \\ \vdots & 0 & 0 & \vdots & \vdots \\ \vdots & \vdots & \cdots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & \mathbf{R}_N \end{pmatrix}.$$

Model uncertainty

In practice models contain errors, due to

- Inaccurate parameter specifications
- Inaccurate parametrisations of sub-grid physical processes
- Inaccurate specification of boundary conditions
- Numerical schemes only approximate solutions
- Poor model resolution

We wish to account for this in the data assimilation process.

Model uncertainty

Assumptions

- Linear model
- Additive model error
- Model error unbiased, Gaussian, random

Model uncertainty

Assumptions

- Linear model
- Additive model error
- Model error unbiased, Gaussian, random

True model

$$\mathbf{x}^t_{i+1} = \mathbf{M}^t_i \mathbf{x}^t_i$$

Model uncertainty

Assumptions

- Linear model
- Additive model error
- Model error unbiased, Gaussian, random

True model

$$\mathbf{x}_{i+1}^t = \mathbf{M}_i^t \mathbf{x}_i^t$$

Erroneous model

$$\mathbf{x}_{i+1}^t = \mathbf{M}_i \mathbf{x}_i^t + \eta_i, \qquad \eta_i \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_i)$$

Modifying the observation error covariance (work with Kat Howes and Alison Fowler)

Consider again the form of the objective function

$$\mathcal{J}(\mathbf{x}_0) = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}^b)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}^b) + \frac{1}{2}(\hat{\mathbf{y}} - \hat{\mathbf{H}}\mathbf{x}_0)^T \mathbf{R}^{*-1}(\hat{\mathbf{y}} - \hat{\mathbf{H}}\mathbf{x}_0),$$

$$\hat{\mathbf{y}} = \begin{pmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \\ \vdots \\ \vdots \\ \vdots \\ \mathbf{y}_N \end{pmatrix} \qquad \hat{\mathbf{H}} = \begin{pmatrix} \mathbf{H}_0 \\ \mathbf{H}_1\mathbf{M}_{0\to 1} \\ \vdots \\ \vdots \\ \vdots \\ \mathbf{H}_N\mathbf{M}_{0\to N} \end{pmatrix}$$

where now $\mathbf{M}$ is the erroneous model and $\mathbf{R}^*$ is the covariance of $\epsilon^*_{ob} = \hat{\mathbf{y}} - \hat{\mathbf{H}}\mathbf{x}^t{}_0$.

Modifying the observation error covariance (work with Kat Howes and Alison Fowler)

Consider again the form of the objective function

$$\mathcal{J}(\mathbf{x}_0) \;=\; \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}^b)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}^b) + \frac{1}{2}(\hat{\mathbf{y}} - \hat{\mathbf{H}}\mathbf{x}_0)^T \mathbf{R}^{*-1}(\hat{\mathbf{y}} - \hat{\mathbf{H}}\mathbf{x}_0),$$

$$\hat{\mathbf{y}} = \begin{pmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \\ \vdots \\ \vdots \\ \vdots \\ \mathbf{y}_N \end{pmatrix} \qquad \hat{\mathbf{H}} = \begin{pmatrix} \mathbf{H}_0 \\ \mathbf{H}_1\mathbf{M}_{0\to 1} \\ \vdots \\ \vdots \\ \vdots \\ \mathbf{H}_N\mathbf{M}_{0\to N} \end{pmatrix}$$

where now $\mathbf{M}$ is the erroneous model and $\mathbf{R}^*$ is the covariance of $\epsilon^*_{ob} = \hat{\mathbf{y}} - \hat{\mathbf{H}}{\mathbf{x}^t}_0$.

How should we specify $\mathbf{R}^*$?

Combined model error and observation error covariance

Let,

$$\mathbf{R}^*{}_{(i,k)} = < \epsilon^*_{ob\,i} (\epsilon^*_{ob\,k})^T > .$$

Then,

$$\mathbf{R}^*{}_{(i,k)} = \begin{cases} \mathbf{R}_0 & \text{for i=k=0} \\ \mathbf{R}_i + \mathbf{H}_i \left[\sum_{j=1}^{min(i,k)} \mathbf{M}_{j\to i} \mathbf{Q}_j \mathbf{M}_{j\to k}{}^T \right] \mathbf{H}_k{}^T & \text{for i=k} \\ \mathbf{H}_i \left[\sum_{j=1}^{min(i,k)} \mathbf{M}_{j\to i} \mathbf{Q}_j \mathbf{M}_{j\to k}{}^T \right] \mathbf{H}_k{}^T & \text{otherwise.} \end{cases} \qquad (1)$$

Combined model error and observation error covariance matrix

$$\mathbf{R}^* = \begin{pmatrix} \mathbf{R}_0 & 0 & \cdots & \cdots & 0 \\ 0 & \mathbf{R}_1 + \mathbf{Q}^*_{(1,1)} & \mathbf{Q}^*_{(1,2)} & \cdots & \mathbf{Q}^*_{(1,N)} \\ \vdots & \mathbf{Q}^*_{(2,1)} & \mathbf{R}_2 + \mathbf{Q}^*_{(2,2)} & \vdots & \vdots \\ \vdots & \vdots & \cdots & \ddots & \vdots \\ 0 & \mathbf{Q}^*_{(N,1)} & \cdots & \cdots & \mathbf{R}_N + \mathbf{Q}^*_{(N,N)} \end{pmatrix}.$$

- increase in block diagonal terms due to model error;
- off diagonal block model error covariance terms (time correlations of model error);
- model error covariance terms vary over time.

Combined model error and observation error covariance matrix

$$\mathbf{R}^* = \begin{pmatrix} \mathbf{R}_0 & 0 & \cdots & \cdots & 0 \\ 0 & \mathbf{R}_1 + \mathbf{Q}^*_{(1,1)} & \mathbf{Q}^*_{(1,2)} & \cdots & \mathbf{Q}^*_{(1,N)} \\ \vdots & \mathbf{Q}^*_{(2,1)} & \mathbf{R}_2 + \mathbf{Q}^*_{(2,2)} & \vdots & \vdots \\ \vdots & \vdots & \cdots & \ddots & \vdots \\ 0 & \mathbf{Q}^*_{(N,1)} & \cdots & \cdots & \mathbf{R}_N + \mathbf{Q}^*_{(N,N)} \end{pmatrix}.$$

- increase in block diagonal terms due to model error;
- off diagonal block model error covariance terms (time correlations of model error);
- model error covariance terms vary over time.

How can we calculate this?

Calculating the modified covariance matrix

Define

$$(\mathbf{d}^o{}_b)_i = \mathbf{y}_i - \mathbf{H}_i \mathbf{M}_{0\to i} \mathbf{x}^b.$$

Calculating the modified covariance matrix

Define

$$(\mathbf{d}^o{}_b)_i = \mathbf{y}_i - \mathbf{H}_i \mathbf{M}_{0\to i} \mathbf{x}^b.$$

Then we can show that

$$E[(\mathbf{d}^o{}_b)_i (\mathbf{d}^o{}_b)_k{}^T] = \mathbf{R}^*{}_{(i,k)} + \mathbf{H}_i \mathbf{M}_{0\to i} \mathbf{B} \mathbf{M}_{0\to k}{}^T \mathbf{H}_k{}^T.$$

Or

$$\mathbf{R}^*{}_{(i,k)} = E[(\mathbf{d}^o{}_b)_i (\mathbf{d}^o{}_b)_k{}^T] - \mathbf{H}_i \mathbf{M}_{0\to i} \mathbf{B} \mathbf{M}_{0\to k}{}^T \mathbf{H}_k{}^T.$$

Idealized coupled nonlinear model

Couples the Lorenz 63 system and 2 linear equations*,

$$\begin{aligned}
\dot{x} &= -\sigma x + \sigma y + \alpha v, \\
\dot{y} &= -xz + rx - y + \alpha w, \\
\dot{z} &= xy - bz, \\
\dot{w} &= -\Omega v - k(w - w^*) - \alpha y, \\
\dot{v} &= \Omega(w - w^*) - kv - \alpha x,
\end{aligned} \tag{2}$$

where $\sigma = 10$, $r = 30$, $b = \frac{8}{3}$, $k = 0.1$, $\Omega = \frac{\pi}{10}$ and $w^* = 2$.

- Runge-Kutta 2nd order method with fixed time step $\Delta t = 0.01$ used to approximate solution of coupled ODE's.
- Consider this as 'true' model.

*F. Molteni *et al.*, (1993), J. Climate.

True idealized coupled nonlinear model

Parameter perturbation method: Stochastic forcing simulation*, but with Gaussian error distributions and random error at each time-step.

The true parameter values σ^t, k^t and α^t change at every time-step,

- ${\sigma_i}^t = \gamma_\sigma \sigma$, where $\gamma_\sigma \sim \mathcal{N}(\mathbf{1}, \frac{1}{12}^2)$,
- ${k_i}^t = \gamma_k k$, where $\gamma_k \sim \mathcal{N}(\mathbf{1}, \frac{1}{6}^2)$,
- ${\alpha_i}^t = \gamma_\alpha \alpha$, where $\gamma_\alpha \sim \mathcal{N}(\mathbf{1}, \frac{1}{12}^2)$,

The difference between the true and erroneous model at each time can be considered as additive model error $\boldsymbol{\eta}_i$ of the form,

$$\mathbf{x}_i^t = \mathcal{M}_{i-1}(\mathbf{x}_{i-1}^t) + \boldsymbol{\eta}_i \qquad i = 1, 2, \ldots 500.$$

* R. Buizza *et al.* (1999), Quart. J. Roy. Met. Soc.

Numerical experiments: design

- Assimilation window length 500 time-steps of length $\Delta t = 0.01$, with all variables observed every 10 time-steps directly. Let $\mathbf{B} = \mathbf{R}_i = 10^{-4}\mathbf{I}$.
- Perturb the true model states using $\mathbf{B}$ and $\mathbf{R}_i$ respectively to produce background model state $\mathbf{x}^b$ and observations $\mathbf{y}_i$.
- Select background vector $\mathbf{x}^b$ and perturb using $\mathbf{B}$ to obtain a sample of 20 background values (note these are all at initial time t_0).
- For each observation time t_i: select observation vector $\mathbf{y}_i$ and perturb using $\mathbf{R}_i$ to obtain a sample of 20 observations.
- Use these samples to estimate $(\mathbf{d}^o{}_b)_i = \mathbf{y}_i - \mathbf{H}_i\mathbf{M}_{0\rightarrow i}\mathbf{x}^b$ at each observation time t_i.
- Take the expectations of the innovation products $E[(\mathbf{d}^o{}_b)_i(\mathbf{d}^o{}_b)_i{}^T]$ at each observation time t_i.
- Calculate $\mathbf{R}^*{}_{(i,i)} = E[(\mathbf{d}^o{}_b)_i(\mathbf{d}^o{}_b)_i{}^T] - \mathbf{H}_i\mathbf{M}_{0\rightarrow i}\mathbf{B}\mathbf{M}_{0\rightarrow i}{}^T\mathbf{H}_i{}^T$.

Compare 4DVar analysis accuracy using $\mathbf{R}^*$ as opposed to $\hat{\mathbf{R}}$.

Numerical experiments: results

- Method 1: use $\hat{\mathbf{R}}$ in 4DVar
- Method 2: use $\mathbf{R}^*$ in 4DVar

Variable	Truth	Analysis Method 1	Error % Method 1	Analysis Method 2	Error % Method 2
x	-3.4866	-3.1111	10.77	-3.4829	0.11
y	-5.7699	-5.2994	8.15	-5.7843	0.25
z	18.341	18.6500	1.68	18.3464	0.03
w	-10.7175	-10.8140	0.90	-10.7181	0.01
v	-7.1902	-7.9787	10.97	-7.1928	0.04

Table : Analysis from Method 1 and Method 2.

Numerical experiments: results

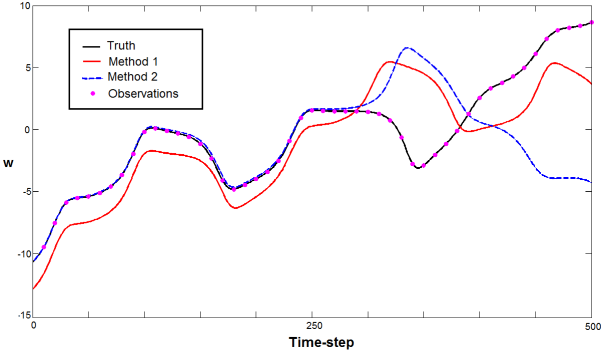

Figure : Trajectory for *w* over the assimilation window.

Weak-constraint 4D-Var (work with Adam El-Said and Nancy Nichols)

Previously we had the objective function

$$\mathcal{J}[\mathbf{x}_0] = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}^b)^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}^b) + \frac{1}{2}\sum_{i=0}^{N}(\mathcal{H}_i[\mathbf{x}_i] - \mathbf{y}_i)^{\mathrm{T}}\mathbf{R}_i^{-1}(\mathcal{H}_i[\mathbf{x}_i] - \mathbf{y}_i)$$

subject to the dynamical system

$$\mathbf{x}_{i+1} = \mathcal{M}_i(\mathbf{x}_i).$$

We now consider the model as a weak constraint

$$\mathbf{x}_{i+1} = \mathcal{M}_i(\mathbf{x}_i) + \boldsymbol{\eta}_i, \qquad \boldsymbol{\eta}_i \sim \mathcal{N}(0, \mathbf{Q}_i)$$

Weak-constraint 4D-Var

Error formulation

$$\mathcal{J}(\mathbf{x}_0, \boldsymbol{\eta}_0, \ldots, \boldsymbol{\eta}_{N-1}) = \mathcal{J}_b + \mathcal{J}_o + \frac{1}{2}\sum_{i=0}^{N-1} \boldsymbol{\eta}_i^T \mathbf{Q}_i^{-1} \boldsymbol{\eta}_i$$

State formulation

$$\begin{aligned} & \mathcal{J}(\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_N) \\ = \; & \mathcal{J}_b + \mathcal{J}_o + \frac{1}{2}\sum_{i=0}^{N-1} (\mathbf{x}_{i+1} - \mathcal{M}_i(\mathbf{x}_i))^T \mathbf{Q}_i^{-1} (\mathbf{x}_{i+1} - \mathcal{M}_i(\mathbf{x}_i)) \end{aligned}$$

Condition number

The expected accuracy of the numerical solution and the speed of convergence are both determined by the condition number of the Hessian.

Condition number

$$\kappa(\mathbf{A}) = ||\mathbf{A}||\,||\mathbf{A}^{-1}||$$

In the matrix 2-norm, for a symm. pos. def. matrix $\mathbf{A}$, we have

$$\kappa(\mathbf{A}) = \lambda_{\text{max}}(\mathbf{A})/\lambda_{\text{min}}(\mathbf{A})$$

Condition number

The expected accuracy of the numerical solution and the speed of convergence are both determined by the condition number of the Hessian.

Condition number

$$\kappa(\mathbf{A}) = ||\mathbf{A}||||\mathbf{A}^{-1}||$$

In the matrix 2-norm, for a symm. pos. def. matrix $\mathbf{A}$, we have

$$\kappa(\mathbf{A}) = \lambda_{\text{max}}(\mathbf{A})/\lambda_{\text{min}}(\mathbf{A})$$

What are the condition numbers of the Hessians of the two formulations sensitive to?
(extends previous work on strong-constraint case by Haben, Lawless and Nichols)

Hessian - Error formulation: $\mathbf{S}_p = \nabla^2 \mathcal{J}(x_0, \eta_i)$

$$\mathbf{S}_p = \mathbf{D}^{-1} + \mathbf{L}^{-T}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{L}^{-1}$$

$$\mathbf{S}_p = \begin{pmatrix} B_0^{-1} & & & \\ & Q_1^{-1} & & \\ & & \ddots & \\ & & & Q_n^{-1} \end{pmatrix} +$$

$$\begin{pmatrix} H_0^T & (H_1M_1)^T & (H_2M_2M_1)^T & \dots & (H_nM_n\dots M_1)^T \\ & H_1^T & (H_2M_2)^T & \dots & (H_nM_n\dots M_2)^T \\ & & H_2^T & \ddots & \vdots \\ & & & \ddots & (H_nM_n)^T \\ & & & & H_n^T \end{pmatrix} \mathbf{R}^{-1} \begin{pmatrix} H_0 & & & & \\ H_1M_1 & H_1 & & & \\ H_2M_2M_1 & H_2M_2 & H_2 & & \\ \vdots & & \ddots & \ddots & \\ H_nM_n\dots M_1 & \dots & & H_nM_n & H_n \end{pmatrix}$$

Hessian - State formulation: $\mathbf{S}_x = \nabla^2 \mathcal{J}(x_i)$

$$\mathbf{S}_x = \mathbf{L}^T \mathbf{D}^{-1} \mathbf{L} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}$$

$$\mathbf{S}_x = \begin{pmatrix} B_0^{-1} + M_1^T Q_1^{-1} M_1 & -M_1^T Q_1^{-1} & & & & \\ -Q_1^{-1} M_1 & Q_1^{-1} + M_2^T Q_2^{-1} M_2 & -M_2^T Q_2^{-1} & & & \\ & -Q_2^{-1} M_2 & \ddots & & & \\ & & & & -M_{n-1}^T Q_{n-1}^{-1} & \\ & & & -Q_{n-1}^{-1} M_{n-1} & Q_{n-1}^{-1} + M_n^T Q_n^{-1} M_n & -M_n^T Q_n^{-1} \\ & & & & -Q_n^{-1} M_n & Q_n^{-1} \end{pmatrix} + \begin{pmatrix} H_0^T R_0^{-1} H_0 & & & \\ & H_1^T R_1^{-1} H_1 & & \\ & & \ddots & \\ & & & H_n^T R_n^{-1} H_n \end{pmatrix}$$

Assumptions

- Background and Model errors are correlated only spatially. Also, the model error covariance matrices are time invariant:

Assumptions

- Background and Model errors are correlated only spatially. Also, the model error covariance matrices are time invariant:
  - $B_0 = \sigma_b^2 C_B$ and $Q_i = \sigma_q^2 C_Q$ for all $i = 1, ..., N$.

Assumptions

- Background and Model errors are correlated only spatially. Also, the model error covariance matrices are time invariant:
  - $B_0 = \sigma_b^2 C_B$ and $Q_i = \sigma_q^2 C_Q$ for all $i = 1, ..., N$.
  - C_B and C_Q are valid auto-covariance matrices. (Symmetric pos-def)

Assumptions

- Background and Model errors are correlated only spatially. Also, the model error covariance matrices are time invariant:
  - $B_0 = \sigma_b^2 C_B$ and $Q_i = \sigma_q^2 C_Q$ for all $i = 1, ..., N$.
  - C_B and C_Q are valid auto-covariance matrices. (Symmetric pos-def)
- Observation errors are uncorrelated $\implies R_i = \sigma_o^2 I$ for all $i = 0, ..., N$.

Assumptions

- Background and Model errors are correlated only spatially. Also, the model error covariance matrices are time invariant:
  - $B_0 = \sigma_b^2 C_B$ and $Q_i = \sigma_q^2 C_Q$ for all $i = 1, ..., N$.
  - C_B and C_Q are valid auto-covariance matrices. (Symmetric pos-def)
- Observation errors are uncorrelated $\implies$ $R_i = \sigma_o^2 I$ for all $i = 0, ..., N$.
- We take p regularly spaced, direct observations at every time step giving $p(N+1)$ observations in total.

Assumptions

- Background and Model errors are correlated only spatially. Also, the model error covariance matrices are time invariant:
  - $B_0 = \sigma_b^2 C_B$ and $Q_i = \sigma_q^2 C_Q$ for all $i = 1, ..., N$.
  - C_B and C_Q are valid auto-covariance matrices. (Symmetric pos-def)
- Observation errors are uncorrelated $\implies$ $R_i = \sigma_o^2 I$ for all $i = 0, ..., N$.
- We take p regularly spaced, direct observations at every time step giving $p(N+1)$ observations in total.
  - Observations available are less than the size of the state, $p(N+1) < n(N+1)$.

Assumptions

- Background and Model errors are correlated only spatially. Also, the model error covariance matrices are time invariant:
  - $B_0 = \sigma_b^2 C_B$ and $Q_i = \sigma_q^2 C_Q$ for all $i = 1, ..., N$.
  - C_B and C_Q are valid auto-covariance matrices. (Symmetric pos-def)
- Observation errors are uncorrelated $\implies R_i = \sigma_o^2 I$ for all $i = 0, ..., N$.
- We take p regularly spaced, direct observations at every time step giving $p(N+1)$ observations in total.
  - Observations available are less than the size of the state, $p(N+1) < n(N+1)$.
  - We observe at the same positions at each time step.

Assumptions

- Background and Model errors are correlated only spatially. Also, the model error covariance matrices are time invariant:
  - $B_0 = \sigma_b^2 C_B$ and $Q_i = \sigma_q^2 C_Q$ for all $i = 1, ..., N$.
  - C_B and C_Q are valid auto-covariance matrices. (Symmetric pos-def)
- Observation errors are uncorrelated $\implies R_i = \sigma_o^2 I$ for all $i = 0, ..., N$.
- We take p regularly spaced, direct observations at every time step giving $p(N+1)$ observations in total.
  - Observations available are less than the size of the state, $p(N+1) < n(N+1)$.
  - We observe at the same positions at each time step.
- The model $\mathbf{M}$ is a *circulant matrix*.

Error formulation - Bounds

Bounds have been derived on the condition number of the Hessian $\mathbf{S}_p$. These bounds indicate:

- Decreasing σ_o^2 (specifying accurate observations) increases $\kappa(\mathbf{S}_p)$.

Error formulation - Bounds

Bounds have been derived on the condition number of the Hessian $\mathbf{S}_p$. These bounds indicate:

- Decreasing σ_o^2 (specifying accurate observations) increases $\kappa(\mathbf{S}_p)$.
- Longer assimilation windows increase $\kappa(\mathbf{S}_p)$

Error formulation - Bounds

Bounds have been derived on the condition number of the Hessian $\mathbf{S}_p$. These bounds indicate:

- Decreasing σ_o^2 (specifying accurate observations) increases $\kappa(\mathbf{S}_p)$.
- Longer assimilation windows increase $\kappa(\mathbf{S}_p)$
- $\kappa(\mathbf{S}_p)$ is linearly influenced by $\kappa(\mathbf{D})$.

Error formulation - Bounds

Bounds have been derived on the condition number of the Hessian $\mathbf{S}_p$. These bounds indicate:

- Decreasing σ_o^2 (specifying accurate observations) increases $\kappa(\mathbf{S}_p)$.
- Longer assimilation windows increase $\kappa(\mathbf{S}_p)$
- $\kappa(\mathbf{S}_p)$ is linearly influenced by $\kappa(\mathbf{D})$.
- As the difference between σ_b^2 and σ_q^2 increases, so does $\kappa(\mathbf{S}_p)$.

Error formulation - Bounds

Bounds have been derived on the condition number of the Hessian $\mathbf{S}_p$. These bounds indicate:

- Decreasing σ_o^2 (specifying accurate observations) increases $\kappa(\mathbf{S}_p)$.
- Longer assimilation windows increase $\kappa(\mathbf{S}_p)$
- $\kappa(\mathbf{S}_p)$ is linearly influenced by $\kappa(\mathbf{D})$.
- As the difference between σ_b^2 and σ_q^2 increases, so does $\kappa(\mathbf{S}_p)$.
- $\kappa(B_0)$ and $\kappa(Q_i)$ increase as correlation length-scales are increased

State formulation - Bounds

Bounds have been derived on the condition number of the Hessian $\mathbf{S}_x$. These bounds indicate:

- As observation density decreases, the sensitivity to length of assimilation window increases.

State formulation - Bounds

Bounds have been derived on the condition number of the Hessian $\mathbf{S}_x$. These bounds indicate:

- As observation density decreases, the sensitivity to length of assimilation window increases.
- Interesting feature: $\kappa(\mathbf{S}_x)$ is *immune* to assimilation window length iff observation operator is *full rank*

State formulation - Bounds

Bounds have been derived on the condition number of the Hessian $\mathbf{S}_x$. These bounds indicate:

- As observation density decreases, the sensitivity to length of assimilation window increases.
- Interesting feature: $\kappa(\mathbf{S}_x)$ is *immune* to assimilation window length iff observation operator is *full rank*
- Otherwise: as assimilation window length increases, so does $\kappa(\mathbf{S}_x)$. Due to $\lambda_{min}(\mathbf{L}^T\mathbf{D}^{-1}\mathbf{L}) \rightarrow 0$, (second derivative matrix).

State formulation - Bounds

Bounds have been derived on the condition number of the Hessian $\mathbf{S}_x$. These bounds indicate:

- As observation density decreases, the sensitivity to length of assimilation window increases.
- Interesting feature: $\kappa(\mathbf{S}_x)$ is *immune* to assimilation window length iff observation operator is *full rank*
- Otherwise: as assimilation window length increases, so does $\kappa(\mathbf{S}_x)$. Due to $\lambda_{min}(\mathbf{L}^T\mathbf{D}^{-1}\mathbf{L}) \rightarrow 0$, (second derivative matrix).
- *Very* sensitive to condition number of $\mathbf{D}$, with the addition of increased sensitivity to σ_q (in comparison to $\mathbf{S}_p$).

Numerical Results: Error formulation

- We demonstrate the bounds by viewing the condition number as a function of correlation length-scale:

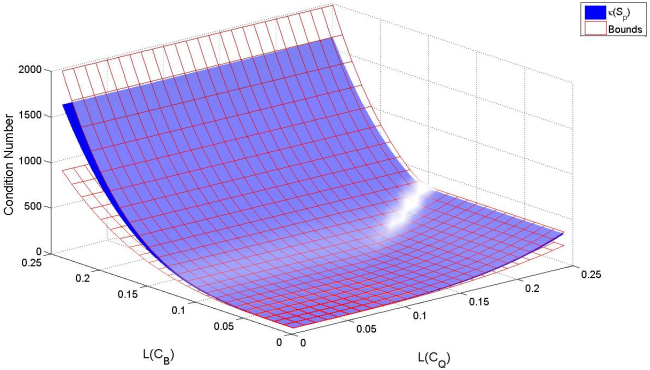

Figure : $\kappa(\mathbf{S}_p)$ (blue-surface) and bounds (red-mesh surfaces) as a function of $L(C_B)$ and $L(C_Q)$.

Numerical Results: State formulation

The $\mathcal{J}(x_i)$ formulation is more sensitive to $\kappa(\mathbf{D})$:

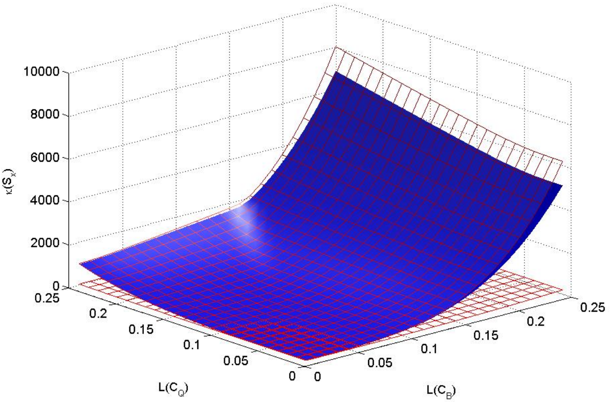

Figure : Surface plot of $\kappa(\mathbf{S}_x)$ (blue surface) and bounds (red mesh). Non-vertical axes measure background error correlation length-scale $L(C_B)$ and

Numerical Results: State formulation

- We now consider the condition number as a function of assimilation window length and observation density:

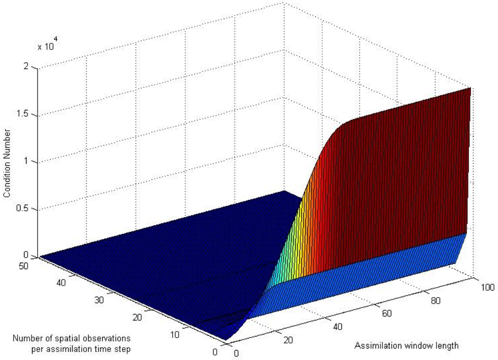

Figure : Surface plot of $\kappa(\mathbf{S}_x)$ (blue surface). Vertical axis measures condition number. The non-vertical axes measure spatial observation

Summary

- Allowing for model uncertainty becoming more important in data assimilation.

Summary

- Allowing for model uncertainty becoming more important in data assimilation.
- Appropriate inflation of covariances can increase accuracy of state estimate at initial time.

Summary

- Allowing for model uncertainty becoming more important in data assimilation.
- Appropriate inflation of covariances can increase accuracy of state estimate at initial time.
  - Does not require prior knowledge of model error statistics.

Summary

- Allowing for model uncertainty becoming more important in data assimilation.
- Appropriate inflation of covariances can increase accuracy of state estimate at initial time.
  - Does not require prior knowledge of model error statistics.
- Weak-constraint formulations a possible way forward - Requires specification of statistics.

Summary

- Allowing for model uncertainty becoming more important in data assimilation.
- Appropriate inflation of covariances can increase accuracy of state estimate at initial time.
  - Does not require prior knowledge of model error statistics.
- Weak-constraint formulations a possible way forward - Requires specification of statistics.
- Different formulations lead to optimization problems with different sensitivities.